Genome variation in the hyperthermophilic archaeon *Aeropyrum*

Takashi Daifuku, Takashi Yoshida, and Yoshihiko Sako*

Graduate School of Agriculture; Kyoto University; Kyoto, Japan

Keywords: hyperthermophilic archaea, genome synteny, virus, CRISPR, pan-genome

Abbreviations: ORF, open reading frame; APSV1, *Aeropyrum pernix* spindle-shaped virus 1; APOV1, *Aeropyrum pernix* ovoid virus 1; CRISPR, clustered regularly interspaced short palindromic repeats

Submitted: 09/12/2013

Revised: 10/15/2013

Accepted: 10/15/2013

Citation: Daifuku T, Yoshida T, Sako Y. Genome variation in the hyperthermophilic archaeon *Aeropyrum*. Mobile Genetic Elements 2013; 3:e26833; <http://dx.doi.org/10.4161/mge.26833>

*Correspondence to: Yoshihiko Sako;
Email: sako@kais.kyoto-u.ac.jp

Commentary to: Daifuku T, Yoshida T, Kitamura T, Kawaichi S, Inoue T, Nomura K, Yoshida Y, Kuno S, Sako Y. Variation of the virus-related elements within syntenic genomes of the hyperthermophilic Archaeon *Aeropyrum*. Appl Environ Microbiol 2013; 79:5891–8; PMID:23872576; <http://dx.doi.org/10.1128/AEM.01089-13>

Aeropyrum spp are aerobic, heterotrophic, and hyperthermophilic marine archaea. There are two closely related *Aeropyrum* species, *Aeropyrum camini* and *Aeropyrum pernix*, which are isolated from geographically distinct locations. Recently, we compared their genome sequences to determine their genomic variation. They possess highly conserved small genomes, reflecting their close relationship. The entire genome similarity may result from their survival strategies in adapting to extreme environmental conditions. Meanwhile, synteny disruptions were observed in some regions including clustered regularly interspaced short palindromic repeats elements. Further, the largest portion of their non-orthologous genes were genes in the two proviral regions of *A. pernix* (*Aeropyrum pernix* spindle-shaped virus 1 and *Aeropyrum pernix* ovoid virus 1) or ORFans considered to be derived from viruses. Our data shows that genomic diversification of *Aeropyrum* spp may be substantially induced by viruses. This suggests that *Aeropyrum* spp may have a large pan-genome that can be extended by viruses, while each of the species shares a highly conserved small genome specializing for extreme environments.

The members of the genus *Aeropyrum* are heterotrophic, aerobic, and hyperthermophilic crenarchaea.^{1,2} The genus is composed of two closely related species, *Aeropyrum camini* and *Aeropyrum pernix*. The type strains are *A. camini*

SY1 and *A. pernix* K1. *A. camini* SY1 was isolated from the surface of a deep-sea hydrothermal vent chimney at the Suiyo Seamount in the Izu-Bonin Arc, Japan, at a depth of 1385 m;¹ and *A. pernix* K1 was isolated from a coastal solfataric thermal vent in Kodakara-Jima Island in southwestern Japan.²

Comparative genome analysis between closely related organisms shows relevant physiological differences depending on their habitat, which are reflected in genome divergence.^{3,4} We recently determined the complete genome sequence of *A. camini* and compared it with the *A. pernix* genome⁵ to determine the genetic differences between close relatives in distinct habitats.⁶

The general genome features were similar between *A. camini* and *A. pernix*: small genome size (1.60 to 1.67 Mbp), G+C content (56.3 to 56.7%), and number of open reading frames (ORFs) (1645 to 1700). The genomic similarity score, which is calculated using the bit scores of shared orthologous genes,⁷ was 0.87 and nucleotide identity between the two chromosomes was 73.2 to 76.6% over a wide range of the chromosomes (86.2 to 90.2%), indicating their close relationship. The genomes also showed high synteny. Within other closely related archaea and bacteria, comparative genomics shows extensive rearrangements in the nucleotide alignment and poorly conserved gene order.⁸

One of the factors promoting genomic synteny is possibly the lack of the RecBCD system, a well-characterized recombinational enzyme complex in

Figure 1. Two proviral regions (APOV1 and APSV1) were present in *A. pernix* and absent in *A. camini*. Proviral regions, tRNAs, and open reading frames (ORFs) are shown as red boxes, green vertical lines, and arrows, respectively. Orthologous genes are shown in navy blue highlighted by orange.

bacteria, or the restricted number of mobile elements like insertion sequences and miniature inverted-repeat transposable elements that can be the target sequences for homologous recombination. In addition, we speculate *Aeropyrum* spp are specialized in their extreme and narrow-range habitat at the genomic level and their specialized genome may contribute to their genomic integrity, because disruptions in synteny can lead to alteration in gene regulation associated with lower fitness for their habitat. This was supported by the estimate that 82 to 84% of the horizontally acquired genes identified in the *Aeropyrum* genomes were derived from thermophiles.

We observed some genomic variations, although the two genomes were highly conserved. Remarkable synteny disruptions were identified in the virus-related elements. In general, integrative genetic elements are often found at the tRNA loci,⁹ where the *A. pernix* genome included two proviral regions (Fig. 1). The replication of these proviruses is induced under oxygen-limiting conditions. They are named *Aeropyrum pernix* spindle-shaped virus 1 (APSV1) and *Aeropyrum pernix* ovoid virus 1 (APOV1).¹⁰ Conversely, the two proviral sequences were absent in the *A. camini* genome despite the synteny of neighboring regions. Instead, *A. camini* encoded two non-orthologous genes, ORF575 (COG1111, ERCC4-like helicase) and ORF576 (hypothetical protein), next to tRNA21-Val (Fig. 1). The genome synteny was also broken at the clustered regularly interspaced

short palindromic repeats (CRISPR) elements. The CRISPR element is an acquired immune system in archaea and bacteria against viruses and plasmids via incorporation of short sequences, called spacers, derived from foreign genetic elements.¹¹ In other words, CRISPR spacers represent signatures of invading genetic elements. The composition of CRISPR spacers was different between *A. camini* and *A. pernix*. When all 144 spacer sequences collected from both genomes were compared against the NCBI nr nucleotide database using BLASTN,¹² three spacers (two spacers in *A. camini* and a spacer in *A. pernix*) and a spacer in *A. camini* significantly matched the genomes of APSV1 and APOV1, respectively, indicating that *A. camini* interacted with viruses closely related to them in the past, and if the CRISPR is functional, *A. camini* may escape viral infection. The other 140 spacers did not show similarity to any other nucleotide sequence in the database. This implies that *Aeropyrum* spp are challenged by diverse and uncharacterized foreign genetic elements in addition to known viruses infecting *A. pernix*.^{13,14}

A detailed analysis of the non-orthologous genes in *Aeropyrum* spp supports our supposition that viruses may be involved in their genome variations. Among the non-orthologous genes, the major portion of them (41 to 45%) were classified into ORFans, which did not show similarity to any other available protein sequence and are potentially derived from viruses, and proviral genes. We conclude that *Aeropyrum* spp interact with diverse viruses and their genomic

diversification may be substantially due to viruses, notwithstanding their conserved genomes specializing in extreme environment.

The variable gene component is responsible for expanding physiological and ecological capabilities of microorganisms,¹⁵ most notably antibiotic resistance among bacterial pathogens.¹⁶ Although the variable genes in *Aeropyrum* were mostly derived from viruses with unknown functions, they are potentially responsible for the acquisition of new functions. If every unique microbial strain contains a different set of variable genes, the size of the pan-genome (the total set of genes found in all strains¹⁷) becomes vast in a large microbial population. Each *Aeropyrum* strain appears to share conserved small genomes encoding genes required for cell maintenance and, at the same time the *Aeropyrum* population's pan-genome may be extended by viruses to give a significant genetic reservoir exploited for adaptive purposes, resulting in the increased fitness of the population in changeable extreme environments.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

This work was supported by Grant-in-Aid for Science Research (no. 20248023) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. T.D. is a research fellow supported by JSPS for Young Scientists (no. 252043).

References

1. Nakagawa S, Takai K, Horikoshi K, Sako Y. *Aeropyrum camini* sp. nov., a strictly aerobic, hyperthermophilic archaeon from a deep-sea hydrothermal vent chimney. *Int J Syst Evol Microbiol* 2004; 54:329-35; PMID:15023940; <http://dx.doi.org/10.1099/ijs.0.02826-0>
2. Sako Y, Nomura N, Uchida A, Ishida Y, Morii H, Koga Y, Hoaki T, Maruyama T. *Aeropyrum pernix* gen. nov., sp. nov., a novel aerobic hyperthermophilic archaeon growing at temperatures up to 100 °C. *Int J Syst Bacteriol* 1996; 46:1070-7; PMID:8863437; <http://dx.doi.org/10.1099/00207713-46-4-1070>
3. Rocap G, Larimer FW, Lamerdin J, Malfatti S, Chain P, Ahlgren NA, Arellano A, Coleman M, Hauser L, Hess WR, et al. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* 2003; 424:1042-7; PMID:12917642; <http://dx.doi.org/10.1038/nature01947>
4. Gunbin KV, Afonnikov DA, Kolchanov NA. Molecular evolution of the hyperthermophilic archaea of the *Pyrococcus* genus: analysis of adaptation to different environmental conditions. *BMC Genomics* 2009; 10:639; PMID:20042074; <http://dx.doi.org/10.1186/1471-2164-10-639>
5. Kwarabayasi Y, Hino Y, Horikawa H, Yamazaki S, Haikawa Y, Jin-no K, Takahashi M, Sekine M, Baba S, Ankai A, et al. Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1. *DNA Res* 1999; 6:83-101, 145-52; PMID:10382966; <http://dx.doi.org/10.1093/dnares/6.2.83>
6. Daifuku T, Yoshida T, Kitamura T, Kawaichi S, Inoue T, Nomura K, Yoshida Y, Kuno S, Sako Y. Variation of the virus-related elements within syntenic genomes of the hyperthermophilic Archaeon *Aeropyrum*. *Appl Environ Microbiol* 2013; 79:5891-8; PMID:23872576; <http://dx.doi.org/10.1128/AEM.01089-13>
7. Alcaraz LD, Moreno-Hagelsieb G, Eguarte LE, Souza V, Herrera-Estrella L, Olmedo G. Understanding the evolutionary relationships and major traits of *Bacillus* through comparative genomics. *BMC Genomics* 2010; 11:332; PMID:20504335; <http://dx.doi.org/10.1186/1471-2164-11-332>
8. Novichkov PS, Wolf YI, Dubchak I, Koonin EV. Trends in prokaryotic evolution revealed by comparison of closely related bacterial and archaeal genomes. *J Bacteriol* 2009; 191:65-73; PMID:18978059; <http://dx.doi.org/10.1128/JB.01237-08>
9. Wozniak RAF, Waldor MK. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev Microbiol* 2010; 8:552-63; PMID:20601965; <http://dx.doi.org/10.1038/nrmicro2382>
10. Mochizuki T, Sako Y, Prangishvili D. Provirus induction in hyperthermophilic archaea: characterization of *Aeropyrum pernix* spindle-shaped virus 1 and *Aeropyrum pernix* ovoid virus 1. *J Bacteriol* 2011; 193:5412-9; PMID:21784945; <http://dx.doi.org/10.1128/JB.05101-11>
11. Sorek R, Kunin V, Hugenholtz P. CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* 2008; 6:181-6; PMID:18157154; <http://dx.doi.org/10.1038/nrmicro1793>
12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215:403-10; PMID:2231712
13. Mochizuki T, Yoshida T, Tanaka R, Forterre P, Sako Y, Prangishvili D. Diversity of viruses of the hyperthermophilic archaeal genus *Aeropyrum*, and isolation of the *Aeropyrum pernix* bacilliform virus 1, APBV1, the first representative of the family *Clavaviridae*. *Virology* 2010; 402:347-54; PMID:20430412; <http://dx.doi.org/10.1016/j.virol.2010.03.046>
14. Mochizuki T, Krupovic M, Pehau-Arnaudet G, Sako Y, Forterre P, Prangishvili D. Archaeal virus with exceptional virion architecture and the largest single-stranded DNA genome. *Proc Natl Acad Sci U S A* 2012; 109:13386-91; PMID:22826255; <http://dx.doi.org/10.1073/pnas.1203668109>
15. Gogarten JP, Townsend JP. Horizontal gene transfer, genome innovation and evolution. *Nat Rev Microbiol* 2005; 3:679-87; PMID:16138096; <http://dx.doi.org/10.1038/nrmicro1204>
16. Dobrindt U, Hochhut B, Hentschel U, Hacker J. Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* 2004; 2:414-24; PMID:15100694; <http://dx.doi.org/10.1038/nrmicro884>
17. Medini D, Donati C, Tettelin H, Maignani V, Rappuoli R. The microbial pan-genome. *Curr Opin Genet Dev* 2005; 15:589-94; PMID:16185861; <http://dx.doi.org/10.1016/j.gde.2005.09.006>